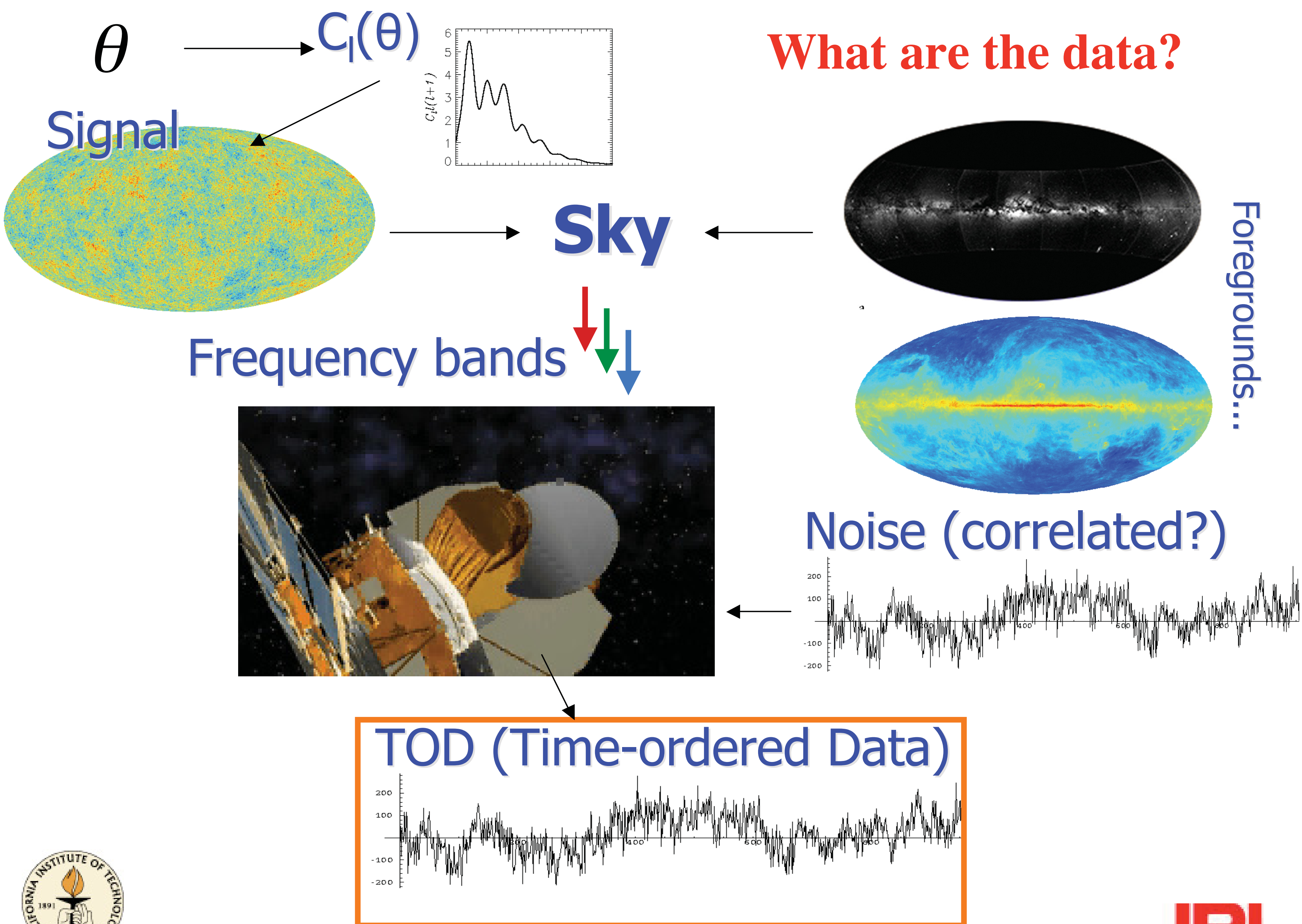
Hybrid Gibbs Sampling and MCMC for CMB Analysis at Small Angular Scales

J. Jewell (JPL)

With H.K. K. Eriksen, B.D. Wandelt, K.M. Gorski, G. Huey, I.J. O'Dwyer, C. Dickinson, A.J. Banday, C.R. Lawrence

What are the data?



Simulation and Inference

Joint density of "everything":

$$p(d,s,\theta) = p(d\,|\,s)\,p(s\,|\,\theta)\,p(\theta)$$

Data Underlying "truth" Model parameters

Simulation: Condition on the model

$$p(d,s\,|\,\theta) = p(d\,|\,s)\,p(s\,|\,\theta)$$

Inference: Condition on the data

$$p(\theta,s\,|\,d) \propto p(d\,|\,s)\,p(s\,|\,\theta)\,p(\theta)$$

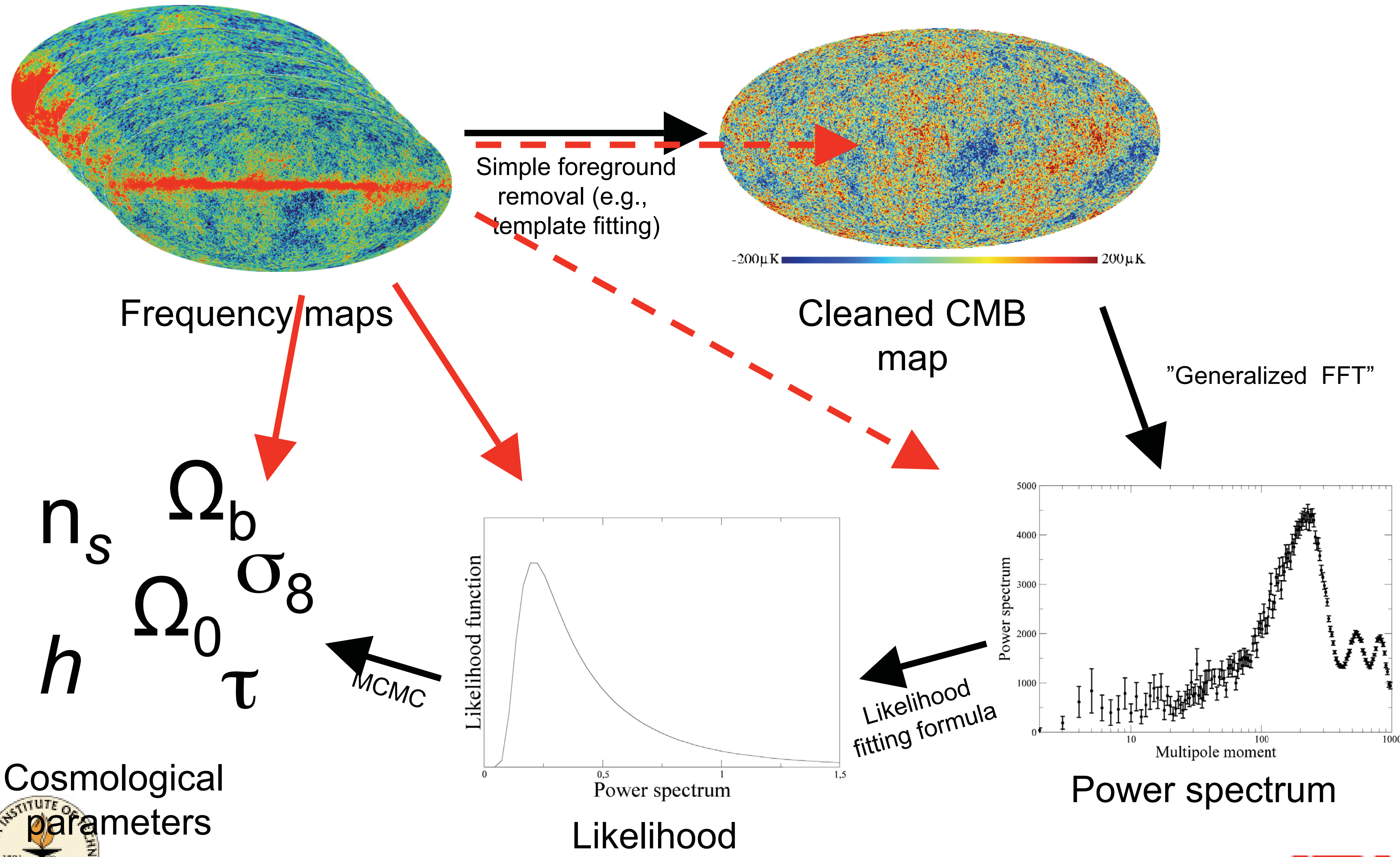
Factors in joint density given CMB data:

$$-2\log p(d\,|\,s) \propto (d_\nu - A_\nu s)N_\nu^{-1}(d_\nu - A_\nu s) = \chi^2$$

$$-2\log p(s\,|\,\theta) = sC^{-1}(\theta)s + \log\left|C(\theta)\right|$$

$$p(\theta) = \text{Prior on parameters}$$

Bayesian CMB Analysis - Can we Beat O(N^3)??



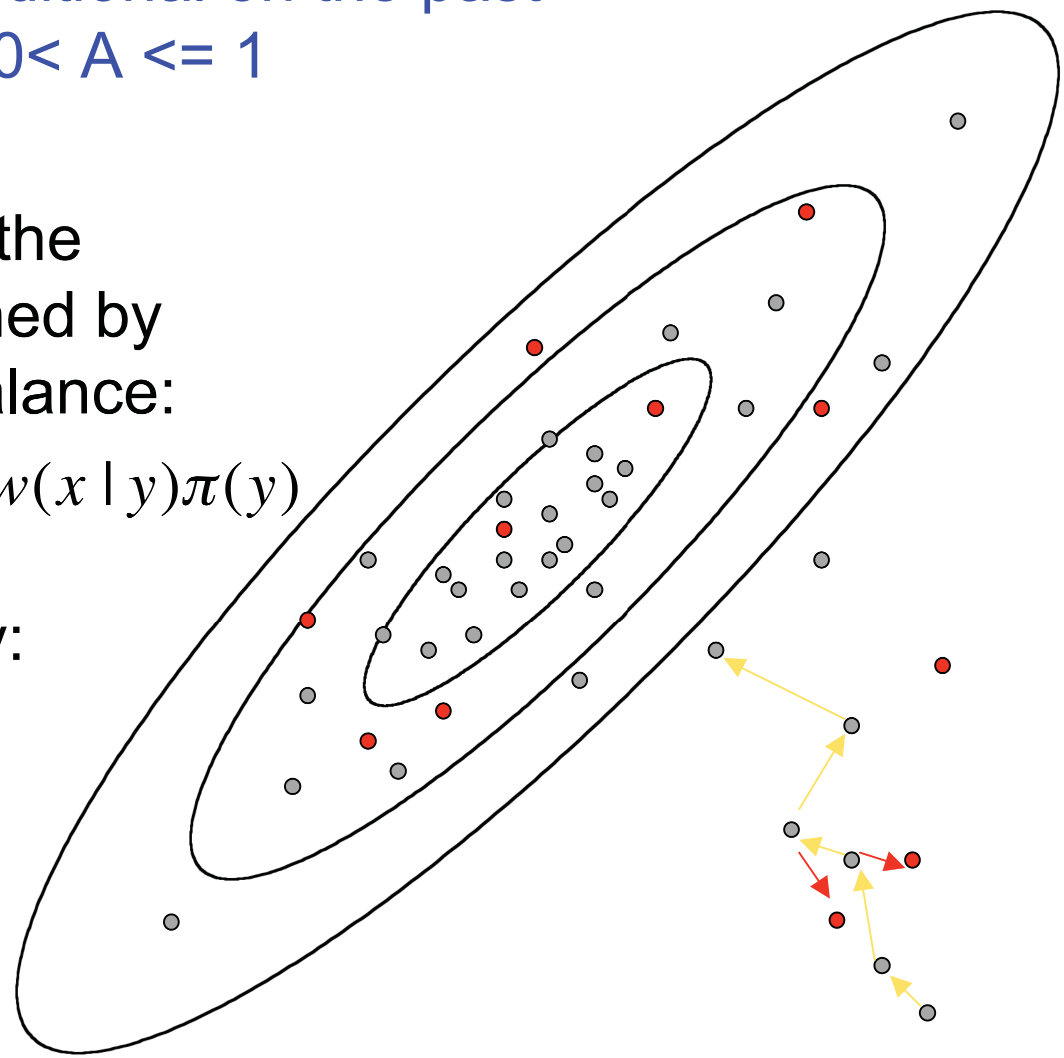
Mapping the Posterior with Metropolis-Hastings MCMC

Algorithm:

1) Propose new state, conditional on the past
2) Accept with probability 0< A <= 1
3) Continue

For any "proposal" matrix, the accept probability determined by the condition of detailed balance:

$$\pi(x)w(y \mid x)A(y \mid x) = A(x \mid y)w(x \mid y)\pi(y)$$

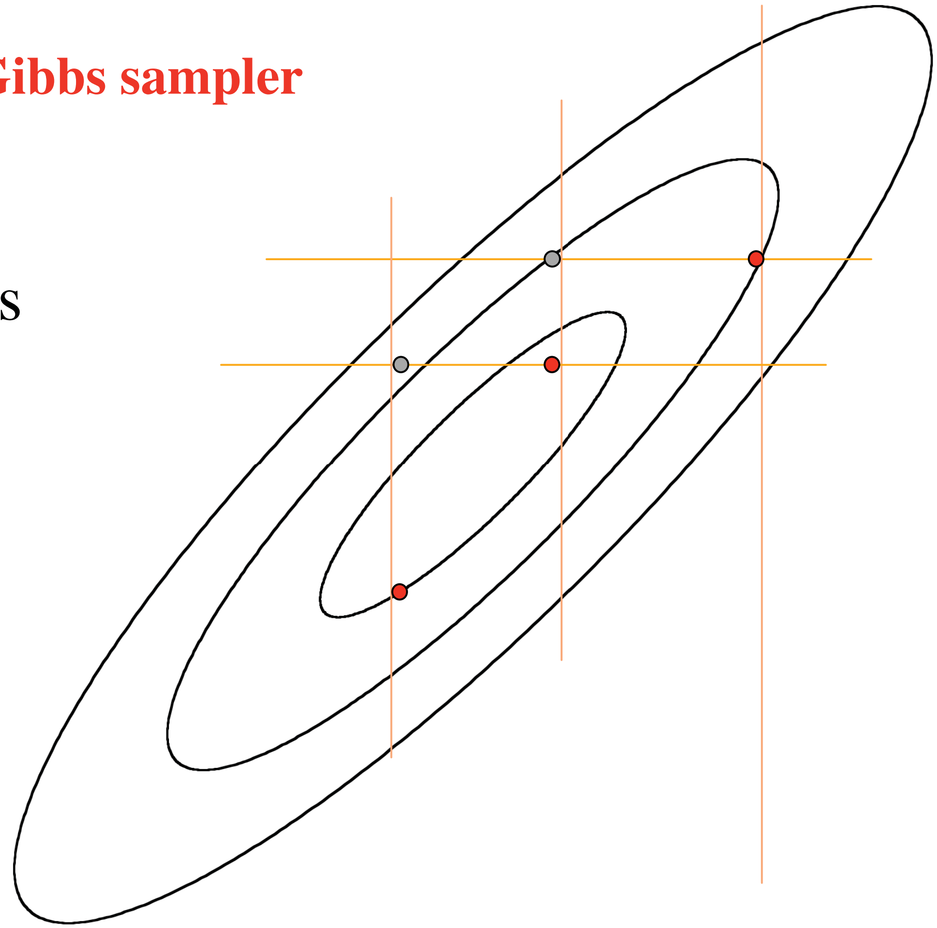


Maximal Accept Probability:

$$A(y \mid x) = \min\left[1, \frac{\pi(y)w(x \mid y)}{\pi(x)w(y \mid x)}\right]$$

Special Case of MH MCMC: The Gibbs sampler

-Sequentially propose variations from conditional densities…
-Accept probability is unity!!



Gibbs Sampling for CMB:

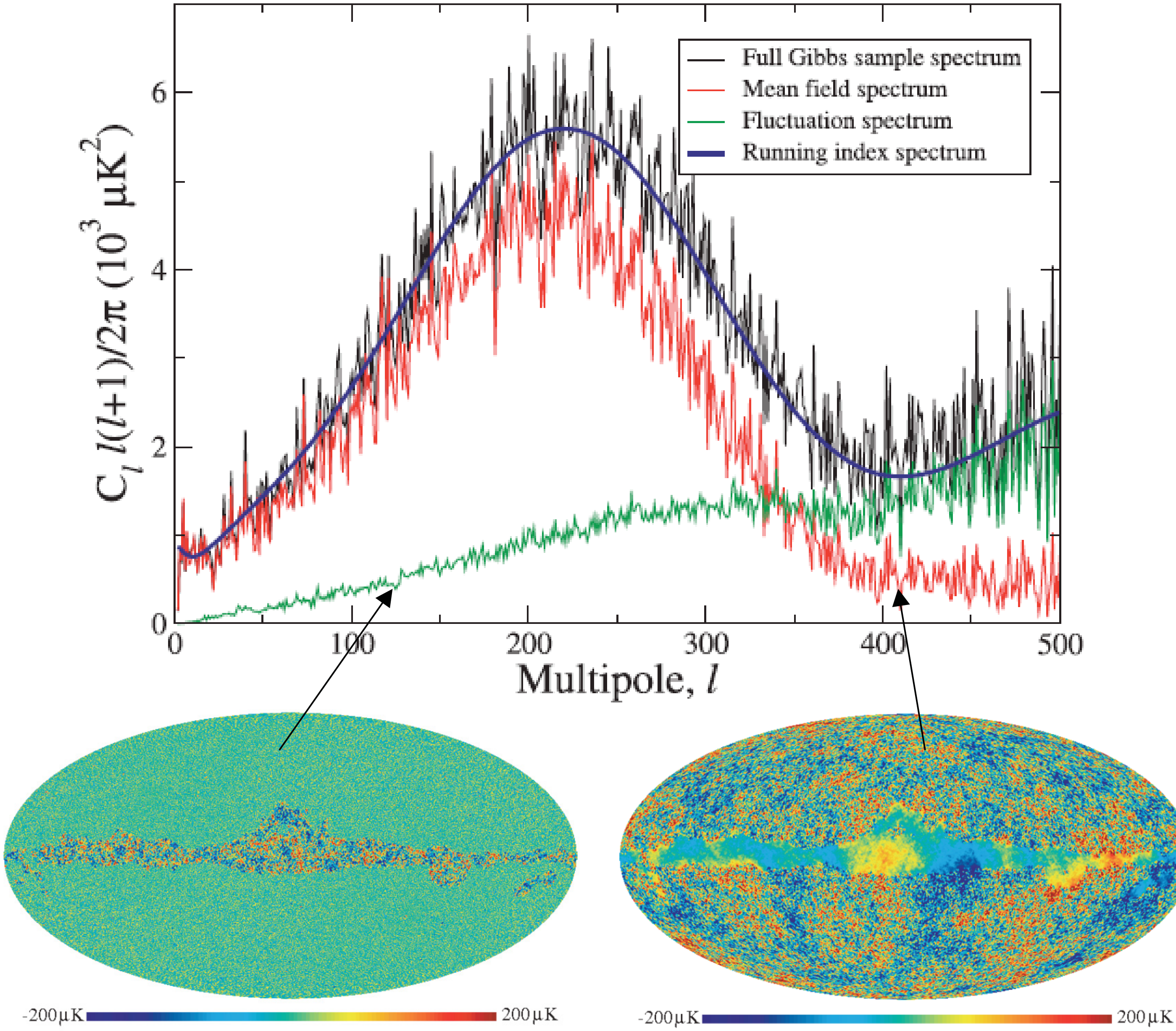
$$s^{(i+1)} \leftarrow p(s \mid C_l^{(i)}, d)$$

$$C_l^{(i+1)} \leftarrow p(C_l \mid s^{(i+1)}, d) = p(C_l \mid s^{(i+1)})$$

Method originally presented in:

- Jewell, et al., ApJ, 609,1,2004
- Wandelt et al., Phys. Rev. D., 70,083511,2004

Sampling the CMB given the Power Spectrum and Data

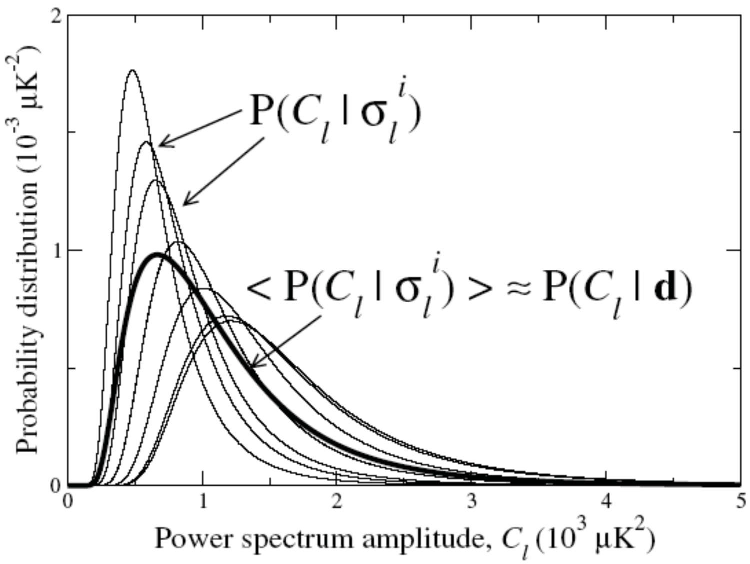


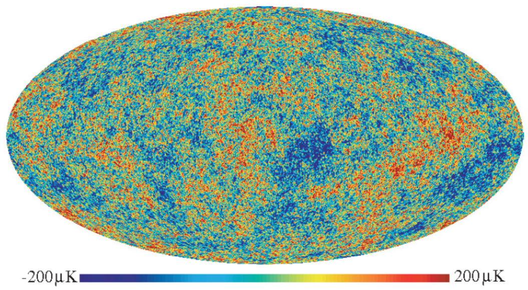
J. Jewell, 2008 Aspen Winter Conference

CMB Gibbs Sampler

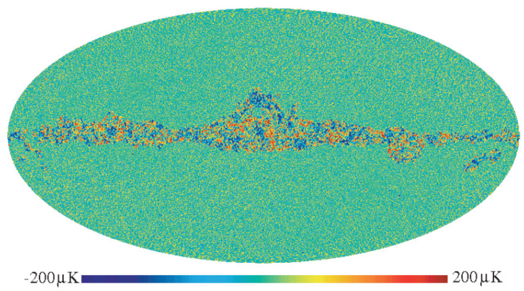
Iterate with: $$p(\theta \mid s)\, p(s \mid d, \theta')$$

$$p(\theta \mid s) \propto p(\theta) \prod_{lm} \frac{e^{-\sigma_l / 2C_l(\theta)}}{\sqrt{2\pi} C_l^{1/2}(\theta)}$$

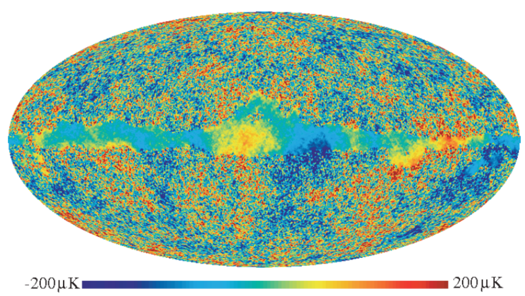




Sum of the two maps is a sample from the conditional



Random variation consistent with our uncertainty



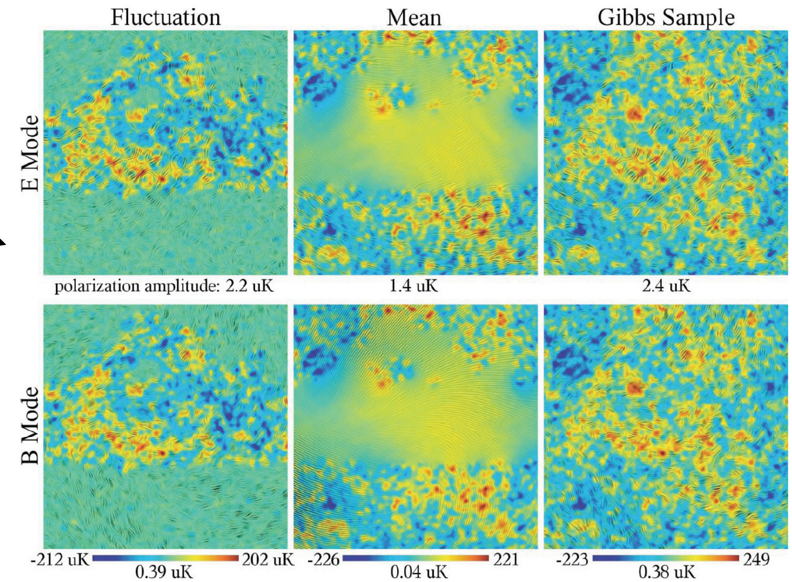
Mean Field map given power spectrum guess

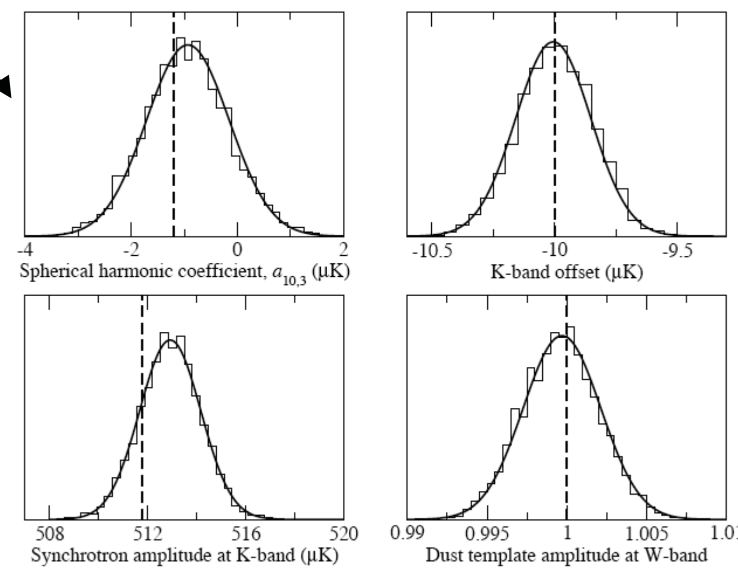
Validation and Applications

- Validation for temperature and polarization:
  - *Power Spectrum Estimation from High-Resolution Maps by Gibbs Sampling*, Eriksen et al., ApJS, 155, 227, 2004
  - *Estimation of Polarized Power Spectra by Gibbs Sampling*, Larson et al., ApJ, 656, 653, 2007
- Extension of method to include foregrounds (temperature data):
  - *Joint Bayesian Component Separation and CMB Power Spectrum Estimation*, Eriksen et al., accepted to ApJ, arXiv 0709.1058
- Applications to WMAP data:
  - *Bayesian Power Spectrum Analysis of the First-Year Wilkinson Microwave Anisotropy Probe Data*, O'Dwyer et al., ApJL, 617, 99, 2004
  - *A Reanalysis of the 3 Year Wilkinson Anisotropy Probe Temperature Power Spectrum and Likelihood*, Eriksen et al., ApJ, 656, 641, 2007
  - *Bayesian Analysis of the Low-Resolution Polarized 3 Year WMAP Sky Maps*, Eriksen et al., ApJL, 665, 1, 2007

- Joint CMB and Foreground analysis of WMAP 3 yr. data:
  - Temperature only - **see Clive Dickinson's talk**, as well as Eriksen et al., ApJL, in press, arXiv 0709.1037
  - Temp. and Polarizaton - **see H.K.K. Eriksen's talk**

- From Gibbs samples to cosmological parameters
  - Chu et al., 2005, Phys. Rev. D., 71, 103002





Comparison of Computational Expense

Direct evaluation: Computational Expense: $O[N^3]$

$$-\log \frac{p(\theta \mid d)}{p(\theta)} = \hat{s}(d)[C(\theta)+N]^{-1}\hat{s}(d) + \log\left\|C(\theta)+N\right\|$$

Gibbs Sampling: Computational Expense: $KO[N^{3/2}]$

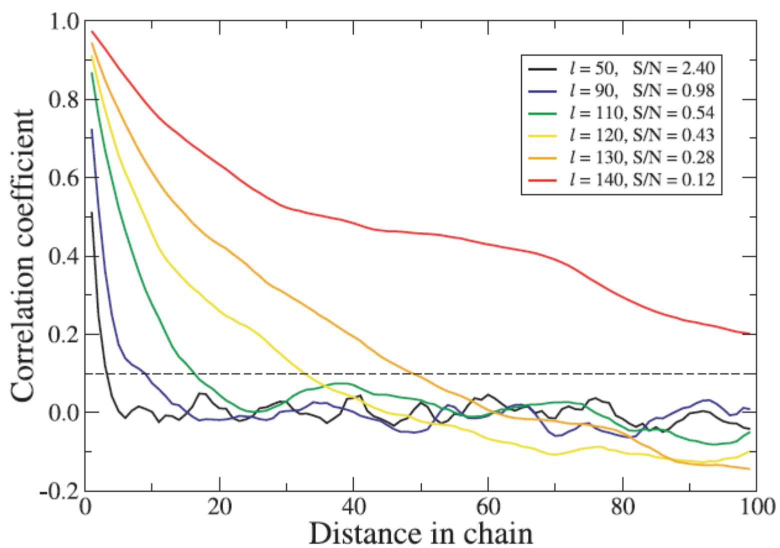
$$\boxed{p(\theta \mid d) \leftarrow_{\infty} \int d\theta' \left[\int ds\; p(\theta \mid s,d)\, p(s \mid \theta',d)\right] p_0(\theta' \mid d)}$$

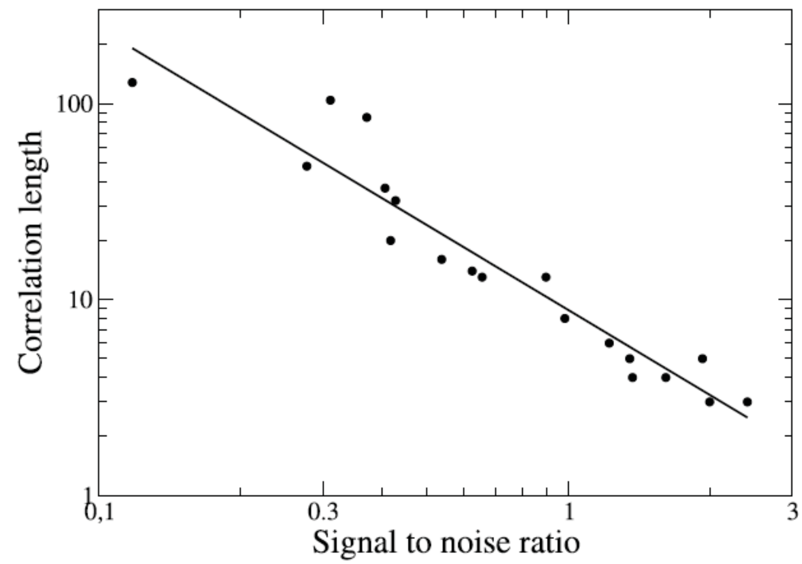
Map-making is the computational bottleneck (scales with expense of mutliplication by N^-1)

Some specific benchmark numbers:

1) Including foregrounds (T only, 5 frequencies = # processors), Nside=64, takes 50 sec/ sample/ freq., or 5000 samples in 350 CPU hours
2) Polarization, Nside=16, dense noise matrix, 2 sec/ sample, or 10^5 samples in 60 CPU hours

Low Signal to Noise, High-L Mixing Properties of Gibbs Sampling





- We want independent samples from joint posterior at all scales
- Gibbs sampling p(C_l | s) at high-L (low S/N) very narrowly peaked (cosmic variance - instead of cosmic AND noise variance)
- Attempting to propose large C_L changes in MCMC, independent of past typically lead to ratio's of matrix determinants which are too expensive to compute…
- Motivates a search for a scheme in which large changes in spectrum can be made with deterministic changes to the CMB map!

Example - Rescale Harmonic Coefficients

"Forward" proposal for CMB map: $s^{(2)} = F(C_l^{(2)}, C_l^{(1)}, s^{(1)}) = (C^{(2)})^{1/2}(C^{(1)})^{-1/2} s^{(1)}$

"Backward" proposal for CMB map: $s^{(1)} = F^{-1}(C_l^{(2)}, C_l^{(1)}, s^{(2)}) = (C^{(1)})^{1/2}(C^{(2)})^{-1/2} s^{(2)}$

Jacobian Factor to be included in Accept Probability:

$$\left|\frac{\partial F}{\partial s^{(2)}}\right|^{-1} = \left|\frac{C^{(2)}}{C^{(1)}}\right|^{1/2}$$

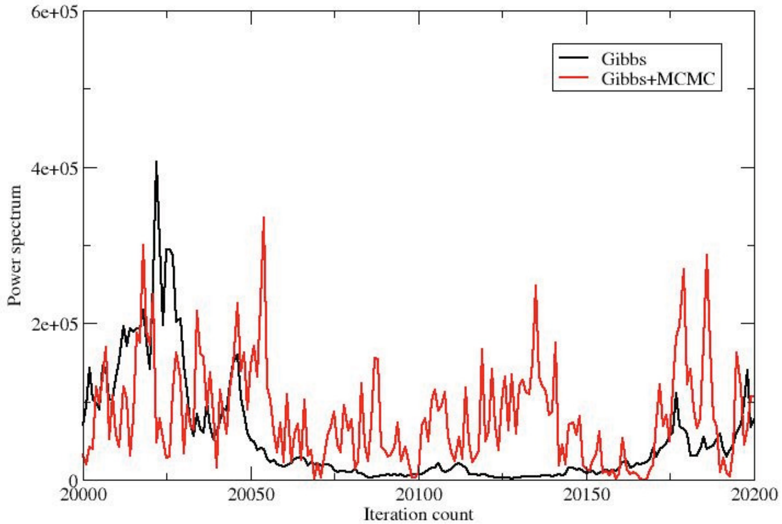
Cancels ratio of determinants in posterior!!

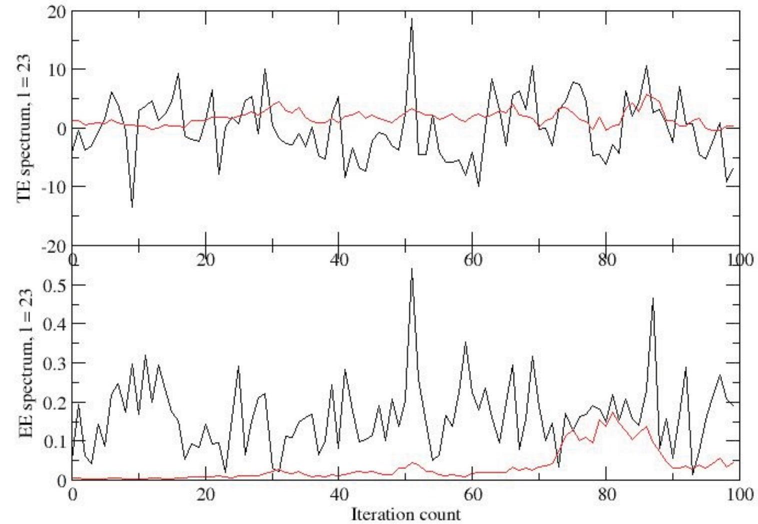
Furthermore - signal "norm" invariant: $s^{(1)}(C^{(1)})^{-1} s^{(1)} = s^{(2)}(C^{(2)})^{-1} s^{(2)}$

$$\boxed{A(C_l^{(2)}, s^{(2)} \mid C_l^{(1)}, s^{(1)}) = \min\left[1, \left(\frac{e^{-(d-s^{(2)})N^{-1}(d-s^{(2)})}}{e^{-(d-s^{(1)})N^{-1}(d-s^{(1)})}}\right) \frac{w(C_l^{(1)} \mid s^{(2)}, C_l^{(2)}, d)}{w(C_l^{(2)} \mid s^{(1)}, C_l^{(1)}, d)}\right]}$$

1. A(1-> 2) depends on change in chi^2!
2. So make large changes to C_L in low S/N regime:
   where standard Gibbs sampling has bad mixing properties!!

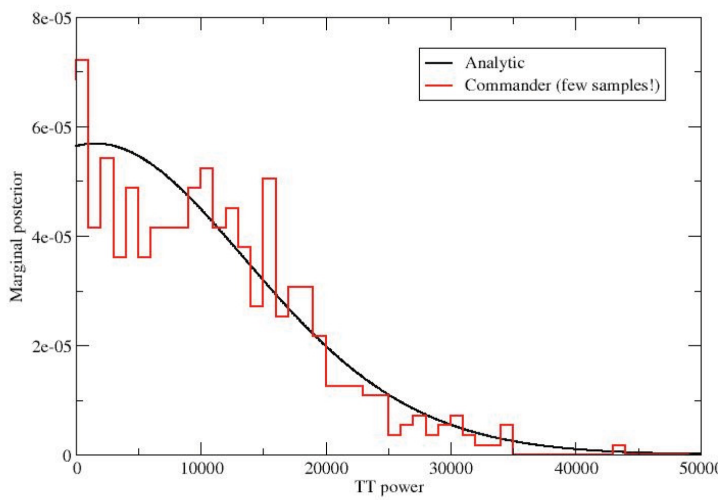
Comparison of Hybrid MCMC+Gibbs to Standard Gibbs





Hybrid MCMC + Gibbs Sampling: Left) Comparison of Gibbs and Gibbs+MCMC power vs. iteration, Right) Comparison for TE and EE power at L=23

New Hybrid MCMC and Gibbs Sampling (L=220 Marginal shown…)



Summary

- Gibbs Sampling has now been validated as an efficient, statistically exact, and practically useful method for "low-L" (as demonstrated on WMAP temperature polarization data)
- *We are extending Gibbs sampling to directly propagate uncertainties in both foreground and instrument models to total uncertainty in cosmological parameters for the entire range of angular scales relevant for Planck*
- Made possible by inclusion of foreground model parameters in Gibbs sampling and hybrid MCMC and Gibbs sampling for the low signal to noise (high-L) regime
- Future items to be included in the Bayesian framework include:
  1. Integration with Hybrid Likelihood (or posterior) code for cosmological parameters
  2. Include other uncertainties in instrumental systematics? (I.e. beam uncertainties, noise estimation, calibration errors, other)